



Libcitations, WorldCat, cultural Impact, and fame

Zuccala, Alesia Ann; Howard D. White

Published in:
Journal of the Association for Information Science and Technology

Publication date:
2018

Citation for published version (APA):
Zuccala, A. A., & Howard D. White (2018). Libcitations, WorldCat, cultural Impact, and fame. *Journal of the Association for Information Science and Technology* , 69(12), 1502-1512.

Libcitations, WorldCat, Cultural Impact, and Fame

Howard D. White
College of Computing and Informatics
Drexel University
Philadelphia, PA 19104 USA
whitehd@drexel.edu

Alesia A. Zuccala
Department of Information Studies
University of Copenhagen
Copenhagen, Denmark-2300
a.zuccala@hum.ku.dk

Abstract

Just as citations to a book can be counted, so can that book's libcitations—the number of libraries in a consortium that hold it. These holdings counts per title can be obtained from the consortium's union catalog, such as OCLC's WorldCat. Librarians seeking to serve their customers well must be attuned to various kinds of merit in books. The result in WorldCat is great variation in the libcitations particular books receive. The higher a title's count (or percentile), the more famous it is—either absolutely or within a subject class. Degree of fame also indicates cultural impact, allowing that further documentation of impact may be needed. Using WorldCat data, we illustrate high, medium, and low degrees of fame with 170 titles published during 1990-1995 or 2001-2006 and spanning the 10 main Dewey classes. We use their total libcitation counts or their counts from members of the Association of Research Libraries, or both, as of late 2011. Our analysis of their fame draws on the recognizability of their authors, the extent to which they and their authors are covered by Wikipedia, and whether they have movie or TV versions. Ordinal scales based on Wikipedia coverage and on libcitation counts are very significantly associated.

Keywords

Librarians, Publishers, Union Catalogs, Percentiles, Wikipedia

Libcitations, WorldCat, Cultural Impact, and Fame

Introduction

Authors want their books to be as widely known as possible. Libcitations have been proposed as an indicator of their success in this regard, as one aspect of cultural impact (White et al., 2009). The present paper illustrates this notion of impact in considerable detail. We show how libcitation counts operationalize it in terms of intuitive notions of fame. Our 170 examples comprise 50 books with the top counts in our sample and 120 books with high, middling, or low counts in the 10 main Dewey classes in two time periods. We also use two independent validation measures. Our counts come, first, from libraries of all types and, second, from the 123 members of the Association for Research Libraries (ARL). The latter are organizations in which evaluative bibliometrics may be of special interest.

Given a particular edition of a book (aka title), libcitations are counts of the number of libraries that hold it. Torres-Salinas & Moed (2009) and Linmans (2010) call the same measure “catalog inclusions.” Plum Analytics (2017) calls it simply “holdings.” The counts are most readily obtained from the union catalogs of sizable cooperatives of libraries. The largest cooperative of this kind is OCLC, whose union catalog WorldCat lists the holdings of more than 16,000 members; public, academic, and research libraries are major types. These libraries are mainly in North America, but they extend around the globe. Books in English dominate their collections, but many also have sizable holdings in other languages. WorldCat may thus reflect popularity on an international scale, with counts ranging from one, in the case of a uniquely held item, to several thousand, in the case of library best-sellers.

Here, we display WorldCat data for a sample of almost 58,000 titles drawn from a research database to be described.

Union catalogs are created by librarians for practical ends such as shared cataloging, cooperative collection development, interlibrary loans, and bibliographic information service. Yet it is possible to view these instruments in an entirely different light—as giant repositories of intelligence about culture. Books preserve accounts of artifacts, sociofacts, and mentifacts in innumerable varieties, and sometimes contribute to cultural change (e.g., *Uncle Tom’s Cabin*, *The Jungle*, *The Other America*). Union catalogs uniquely concentrate information about large numbers of books. Their subject indexing, for instance, can be analyzed for cultural content (e.g., Adler, 2009). Libcitations, by contrast, are a form of indexing that is content-neutral. Taking libraries as paradigmatic institutions of cultural

memory, ranked libcitation counts show that library cooperatives “remember” books to very different extents. Since libraries are both deep-rooted and mappable, these recorded mentions operationalize long-term cultural impact in geographic areas.

Variable cultural impact is thus akin to degrees of fame, which likewise can be measured by counting someone’s or something’s recorded mentions. Canonical works in literatures, for example, achieve fame by being written and talked about in many contexts, and it would be shocking if a union catalog such as WorldCat did not reveal canonical works to be held by numerous libraries in multiple editions. The thousand most widely-held books in WorldCat (OCLC, 2004) are all universal classics (e.g., *Mother Goose*) or highly recognizable reference works. But the fame of contemporary books varies greatly, depending on, e.g., the language in which they are written, their intrinsic subject appeal, the cachet of their publishers (Zuccala et al., 2015), the markets in which they compete, and the publicity and reviews they receive (Zuccala, Someren, & Bellen, 2014; Kousha & Thelwall, 2015; Zhou et al., 2016). Libcitation counts absorb qualitative variables like these by being sales figures. The sales, moreover, are driven not simply by librarians. They rest on chains of judgments by “authors, agents, past editors who have built publishers’ reputations, present-day editors of various kinds, referee-readers, marketers, and wholesalers” (Zuccala & White, 2015: 316). After publication, the chain continues through the contributions of reviewers, other authors (including citers), teachers, media figures, prize-givers, fans, and detractors.

Publication and fame thus intertwine. Publishers bring out copies of books while it serves their interests. This generates the reputations these titles achieve in the short run, and libraries are part of the process. In the long run, libraries make copies of the same titles available after publishers no longer distribute them. Many books are acquired even though there is no explicit demand for them; they are seen as cultural manifestations that deserve at least a modicum of publicity in catalogs and collections. Relatively few books are by household names, of course, but the remaining multitude are at least discoverable parts of the culture; obscurity is not oblivion. Libcitations are thus “heightened” mentions indicating long-term availability to be read and discussed.

Numerous libcitations may not signal a book of high quality. They frequently do, but they may also elevate books that fit various definitions of junk. It is therefore worth noting that even pieces of junk reveal something of their cultural moment and should never be entirely purged. Moreover, millions of WorldCat titles have intellectual value despite low to modest counts. This is especially true of scholarly titles, which by their very nature appeal to limited readerships. Even they may achieve distinction when ranked by their libcitation counts in an appropriate

subject class—that is, when compared to other books in roughly the same subject specialty.

Related Studies

The decision to cite a book and the decision to acquire it leave parallel bibliographic records. The coinage “libcitation”—first syllable as in *library*—underscores the parallelism (Torres-Salinas & Moed, 2009: 11; Linmans, 2010: 339). Since holdings counts have multiple uses in managing and evaluating library collections (White, 2008; Denton 2012), “libcitation” denotes them *as an altmetric for authorial impact*, paralleling the bibliometric use of citations. The citation count and libcitation count of a book are alike in that both usually take considerable time to develop. As a measure of a book’s *reach* or *diffusion*, counting the librarians who have collected it is analogous to counting the citers who have cited it (cf. Ajiferuke & Wolfram, 2010).

However, the motives of citers and librarians are not identical; citers are buttressing claims, while librarians are anticipating interests. As a result, citations and libcitations are not necessarily correlated. (Zuccala & White, 2015, and Kousha & Thelwall, 2016, found correlations that were statistically significant but low.) It is true that some books are both heavily cited and widely held, but other widely-held books have citation records that can only be called unimpressive. They may have *no* citations, or the citations they do have may not appear in the standard indexes—facts not lost on book-oriented humanists and social scientists, who tend to be less well served by standard bibliometric evaluations than journal-oriented scientists (Ochsner, Hug, & Gallerson, 2017). It was with book people in mind that libcitations were proposed as an alternative or complement to citation-based indicators.

In recent years both the Web of Science and Scopus have expanded their coverage of citations to books, and Google Scholar has dramatically improved the situation through its automatic extraction of citation data for all kinds of publications from all corners of the Web (Kousha, Thelwall, & Rezaie, 2011; Gorraiz, Purnell, & Glänzel, 2013). Even so, libcitations may attest to authorial achievements in ways that citations and other altmetrics do not. For instance, in Halevi, Nicholas, & Bar-Ilan (2016) they were available for almost the entire sample of 70,000+ library titles—far more than any of eight other measures. The same is true of libcitations versus other measures in large samples of titles used by Kousha & Thelwall (2015, 2016); Kousha, Thelwall, & Abdoli (2016); and Torres-Salinas, Gumpenberger, & Gorraiz (2017).

Libcitations are nevertheless sometimes dismissed because librarians buy books on automatic pilot; that is, authors accrue libcitations passively, without any particular merit. For instance, Hammarfelt (2016: 122) writes: “Libraries do not always make informed judgments when buying books; they often buy bundles of books.” But

librarians do not acquire bundles of books indiscriminately. Rather, knowing the book world, their budgets, and their actual and potential customers, they arrange to buy some bundles and not others; important criteria include subject matter, language, genre, and publishers' reputations. Considering today's vast output of publications, acquisition of bundles (through, e.g., approval plans or pre-assembled collections) has long been the norm in large libraries and cooperative library systems. Extensive title-by-title selection of books is not feasible (although patron-driven acquisition of individual e-books is increasingly seen). While such buying may be thought to produce uniform collections, we will show that titles are powerfully differentiated in WorldCat by the number of libraries that libcite them. Future studies may relate libcitation counts to types of library acquisitions in more detail.

Some of the framework of citation analysis can be taken over into libcitation analysis. That includes the idea that *sources* can be ranked by the number of *items* they yield. So, for example, if the books in a large WorldCat subject class are ranked by how many libraries each book "yields"—that is, by how many libraries have acquired and cataloged that title—the distribution of counts will exhibit the core-and-scatter structure typical of bibliometrics. Given this structure, libcitation counts for individual scholars or academic departments can be field-normalized or assigned to percentiles just as citations are (Waltman, 2016). Using one variety of normalization, White et al. (2009) evaluated titles by members of the philosophy, history, and political science departments at two Australian universities, and uncovered differences between the matched faculties. The union catalog supplying the libcitation counts was in that case Libraries Australia. The present paper uses a bit of percentile analysis illustratively.

White et al. (2009: 1094) also raised the question whether libcitation counts for books are correlated with their circulation counts. Unfortunately, circulation data for printed books are seldom readily obtainable, and, in any case, librarians buy books not in terms of predicted checkout rates but in terms of what their customers might reasonably expect them to possess; the extent of their circulation is a separate issue. Thus it confuses matters if libcitation counts are interpreted simply as flawed proxies for circulation counts, as in Thelwall (2017: 38):

The number of libraries holding a copy of a book seems to be a reasonable indicator of its likely readership (Torres-Salinas and Moed, 2009; White et al., 2009). It is imperfect because a popular novel might be continually checked out, with a long waiting list, and a course book might be in a university short loan collection so that a different person can check it out every day but other books might never be opened. Similarly, some books are marketed solely as reference works for libraries whereas others are primarily written for the general public.

This implies that libcitation counts can mislead because, in one time and place, (1) the circulation of a widely held novel may be artificially low, (2) the circulation of a course book not widely held may be artificially high, and (3) other books, whatever their libcitation counts, may go unused. It further implies that (4) reference works never

circulate anywhere, and so cannot be compared with (5) books for the general public that circulate everywhere.

With respect to (1) through (3), what happens in one library does not necessarily happen in another. If there were a “WorldCirc” counterpart to WorldCat, it would aggregate circulation counts for books from multiple libraries, and these aggregated counts could tell stories quite different from Thelwall’s. The popular novel, for instance, would benefit from the multi-library counts, especially those from public libraries, which differ from academic libraries in their circulation policies. With respect to (4) and (5), the circulation status of books has little to do with what lib citations actually show, which is *the extent of acquisition across libraries*. What if the popular novel Thelwall mentions is made into a movie? Its high lib citation count would predict this non-scholarly event regardless of its circulation record. Or suppose an author produces a reference work that is bought by a thousand libraries. Would the fact that it never circulates diminish the author’s achievement?

The point of these counter-examples is not to decouple lib citation and circulation entirely. It is to say that neither measure really conveys what goes on in readers’ heads or the uses to which they put what they read. Both merely suggest degrees of impact that need further interpretive comment if they are to be properly judged. A similar point about scholarly and scientific writings is made on the Altmetric (2017) website: “To get at true evidence of impact, you need to dig deeper into the numbers and look at the qualitative data underneath: who’s saying what about research, where in the world research is being cited, reused, read, and so on.” We briefly return to this matter in our conclusion.

Methods

The sample. A version of WorldCat is available free on the Web, but for serious lib citation research, access to the FirstSearch version of WorldCat through an OCLC-member library will be needed. The present paper says little about data-gathering from FirstSearch, however, because the books studied here were not drawn from it. They are a large subset of a sample drawn from Scopus in 2011 (Zuccala & Guns, 2013). That sample consisted of any item cited at least once in journals in Scopus’s *History* or *Literature & literary theory* categories during the windows 1996-2000 or 2007-2011. From this collection of items, *books* and their citation counts were extracted. Although these books were published over a great range of years, the two main groups appeared during 1990-1995 or 2001-2006, the two six-year periods preceding the two citation windows.

One of the Zuccala-Guns questions was whether Scopus citations to books correlated with the same books’ lib citations. In 2012, an OCLC analyst matched the ISBNs of the Scopus books in WorldCat so as to add their

<i>1990-95</i>	<i>2001-06</i>	<i>Main Dewey Class</i>	<i>1990-95</i>	<i>2001-06</i>	<i>Totals</i>	<i>1990-95</i>	<i>2001-06</i>
2%	2%	0 Generalia, computing	308	713	1021	30%	70%
5%	4%	1 Philosophy, psychology	843	1757	2600	32%	68%
9%	9%	2 Religion	1486	3616	5102	29%	71%
40%	36%	3 Social sciences	6874	14589	21463	32%	68%
1%	1%	4 Language	217	563	780	28%	72%
5%	3%	5 Science	783	1381	2164	36%	64%
4%	4%	6 Technology	728	1791	2519	29%	71%
5%	8%	7 Arts, recreation	909	3364	4273	21%	79%
8%	10%	8 Literature	1326	3955	5281	25%	75%
21%	22%	9 History, geography	3645	9047	12692	29%	71%
17119	40776	<i>Totals</i>	17119	40776	57895	30%	70%

Table 1. Distributions of books in the sample when subdivided into main Dewey classes and periods of publication.

libcitation counts as of late 2011. The analyst also added how many members of the Association for Research Libraries held each title. Thus, all books in the Zuccala-Guns database are held by at least one ARL member and at least one non-member. (WorldCat FirstSearch displays only *total* libcitation counts for books, not ARL counts.) In 2013, an OCLC analyst further augmented the file with the books' main Dewey classes, specific Dewey numbers, and other metadata. While main Dewey classes are too coarse to use in judging the libcitations of most individuals, teams, or organizations, they are good for conveying a variety of subject areas compactly.

For the sample here, the Zuccala-Guns database was first reduced to 70,620 unique titles by cutting duplicates. Then another 12,725 titles were cut because they lacked a publication year (36.8%) or Dewey class (32.6%), or were not published in our time periods (30.4%), or were in vestigial classes such as children's books (.02%).

Table 1 has the resulting counts, with a total N of 57,895. The two leftmost columns present the counts of books percentaged to their *column* totals in the two publication periods. The Dewey-class profiles that emerge are very similar and suggest humanities scholars' emphases across subject areas. The two rightmost columns present the raw counts percentaged to the *row* totals. Overall, 30% of the books in the sample were published during 1990-1995, and 70% during 2001-2006. The 30/70 split is roughly duplicated across all Dewey classes. It occurs because Scopus covered many more *History* and *Literature* journals in 2007-2011 than in the earlier period, and so the sample from the 2007-2011 reflects more articles citing more books.

Degrees of fame. As a check on fame independent of libcitations, we used information from the early-2017 Wikipedia. (Searching the English-language edition automatically searches the non-English editions when

appropriate). As a second check, we used whether the title has a movie or TV version in the early-2017 Internet Movie Database. The Wikipedia and IMDB data were then combined to give each title a score on an ordinal scale of fame for our tables.

If a title has its own article in Wikipedia, a separate Wikipedia article almost always exists for its author. Books passing both these tests may also have a movie or TV version; if so, their fame score is 4. Books with both title and author articles are scored 3. A 2 goes to books that lack articles of their own but whose authors have one. (We usually trace only the first author in collaborations.) Books scored 1 have neither author nor title articles, but are cited in Wikipedia at least once (Kousha & Thelwall, 2017). For books scored 0, no information was found.

We expect that many readers will be able to corroborate degrees of fame themselves. That is, they will recognize more books with high libcitation counts than books with lower counts. Such recognition of course requires a particular cultural background. The prototypical OCLC libraries are American and their customers are predominantly English-speaking and U.S.-centered, which strongly affects the libcitations that titles receive. This concession in no way diminishes OCLC's achievement in covering the literatures of many nations and languages. It simply takes into account that *any* union catalog will reflect the dominant characteristics and geographic distribution of its customers, making some titles central and others peripheral.

Results

Neither the Zuccala-Guns database nor our sample was initially created with the present tests of fame in mind, and so our findings simply reveal content that emerged. The first test is blunt: what titles do libcitations bring to the top of our sample regardless of publication year or Dewey class? Table 2 shows the top 20 non-fiction works, all with at least 3,500 libcitations; and the top 10 novels, all with at least 3,100. Most have more than 100 ARL libcitations as well. They have all been widely discussed in various media, and we believe that many of them, or at least their authors, will be instantly recognizable. If not, it is still easy to recognize historical persons and events in their titles that any literate reader would deem significant.

Receipt of prizes has been proposed as one criterion of cultural impact for books (White et al. 2009: 1086), and the authors in Table 2 include winners of the Pulitzer Prize, the National Book Award, and the Nobel Prize, to name only those. In any case, all 30 authors have write-ups in Wikipedia, as do 28 of the books. Fifteen of the 30 have movie or TV versions. These books were intended for wide audiences, and they fulfilled that intent. Table 2 thus sets a standard with which to compare other titles and authors.

<i>Year</i>	<i>Fame</i>	<i>Author</i>	<i>Nonfiction Titles</i>	<i>Libcites</i>	<i>ARLcites</i>
1992	4	McCullough	Truman	4724	137
2001	4	McCullough	John Adams	4655	140
2005	3	McCullough	1776	4429	124
1995	2	C. Powell	My American journey	4233	124
2006	3	Collins	Good to great: why some companies make the leap...and others don't	4195	141
2006	3	Obama	The audacity of hope: thoughts on reclaiming the American dream	4123	125
2005	3	Gladwell	Blink: the power of thinking without thinking	4059	145
1990	4	Branch	Parting the waters: America in the King years, 1954-1963	4039	149
2006	4	Levitt	Freakonomics: a rogue economist explores the hidden side of everything	3967	152
2003	3	H. Clinton	Living history	3949	129
2004	3	Woodward	Plan of attack	3944	136
1993	3	Bennett	The book of virtues: a treasury of great moral stories	3943	80
2005	3	Friedman	The world is flat: a brief history of the twenty-first century	3919	147
2006	3	Goleman	Emotional intelligence: why it can matter more than IQ	3915	134
2005	4	Diamond	Collapse: how societies choose to fail or succeed	3872	147
2005	4	Goodwin	Team of rivals: the political genius of Abraham Lincoln	3840	133
1990	4	Ward / Burns	The Civil War: an illustrated history	3831	107
2001	3	Ambrose	Nothing like it in the world: the men who built the transcontinental railroad 1863-1869	3769	119
1992	3	Gore	Earth in the balance: ecology and the human spirit	3625	149
2006	4	Gore	An inconvenient truth: the planetary emergency of global warming and what we can do about it	3568	132
<i>Year</i>		<i>Author</i>	<i>Fiction Titles</i>	<i>Libcites</i>	<i>ARLcites</i>
2003	4	D. Brown	The Da Vinci code	4725	122
2005	4	Rowling	Harry Potter and the half-blood prince	4507	109
2003	4	Frazier	Cold mountain	4164	106
2001	4	Franzen	The corrections	3676	116
2002	4	Tyler	The accidental tourist	3514	110
1992	3	Michener	Mexico	3324	95
2005	4	Kidd	The mermaid chair	3239	72
2003	3	Morrison	Love	3222	121
2006	4	C. McCarthy	The road	3206	114
2002	3	Crichton	Prey	3151	91

Table 2. Top 20 non-fiction titles and top 10 novels in the entire sample by libcitation counts, with ARL lib citations also shown. Fame scale: 4, Wikipedia title and author entries, plus movie or TV version; 3, Wikipedia title and author entries; 2, Wikipedia author entry.

Table 3 exhibits the 20 titles with the highest ARL lib citation counts. (ARL campuses with more than one holding library cause some counts to exceed 123.) Because these books are more scholarly, they have lower *total* lib citation counts than the books in Table 2, but they still have sold well. Their relatively academic nature is evident in that only three titles have their own Wikipedia articles. However, 16 of the 20 authors have articles of their own; ARL lib citations pick out such renowned American intellectuals as Stanley Fish, Patrick Moynihan, John Rawls, and Arthur Schlesinger. An additional 14 of the books are cited in Wikipedia. Corynne McSherry is cited 22 times there, but not the particular book of hers in the table.

<i>Year</i>	<i>Fame</i>	<i>Author</i>	<i>Nonfiction Titles</i>	<i>ARLcites</i>	<i>Libcites</i>
1994	2	Fish	There's no such thing as free speech, and it's a good thing, too	178	1466
1990	2	Moynihan	On the law of nations	176	1348
1990	2	Rhode, ed.	Theoretical perspectives on sexual difference	176	1136
1990	1	Walker	In defense of American liberties: a history of the ACLU	172	2162
1994	0	Keast	Faculty tenure; a report and recommendations	172	1349
1992	2	Hall	The Oxford companion to the Supreme Court of the United States	171	2790
2001	2	Rawls	Justice as fairness: a restatement	171	1301
2001	2	Patterson	Brown v Board of Education: a civil rights milestone and its troubled legacy	170	2068
1990	2	Thomson	The realm of rights	170	907
1992	2	Annas, ed.	The Nazi doctors and the Nuremberg code: human rights and human experimentation	170	888
2004	3	Schlesinger	The imperial presidency	168	3087
2004	3	Lessig	Free culture: how big media uses technology and the law to lock down culture and control creativity	168	2002
1991	1	Hoff	Law, gender, and injustice: a legal history of U.S. women	168	980
1990	2	Williams	The American Indian in western legal thought: the discourses of conquest	168	878
1992	2	Annas	Gene mapping: using law and ethics as guides	168	775
1994	3	S. Carter	The culture of disbelief: how American law and politics trivialize religious devotion	167	2357
1993	2	Bork	The antitrust paradox: a policy at war with itself	167	1300
1993	2	Massey	American apartheid: segregation and the making of the underclass	166	1538
2001	0	McSherry	Who owns academic work? battling for control of intellectual property	165	944
1991	2	Glendon	Rights talk: the impoverishment of political discourse	165	923

Table 3. Top 20 non-fiction titles in the entire sample by ARL libcitation counts, with total lib citations also shown. Fame scale: 3, Wikipedia title and author entries; 2, Wikipedia author entry; 1, Wikipedia citation; 0, no Wikipedia data.

While the main Dewey class of most titles in Table 3 is “Social sciences,” their *specific* Dewey classes vary widely. Yet together these titles evoke a broad concern that motivated ARL collection developers during our time periods—a concern with matters of law, rights, liberties, justice, ethics, fairness, moral conflict, and governmental policy, especially in the U.S. This hints at the mineable cultural content in WorldCat that was mentioned earlier.

Table 4 shows how many total lib citations and ARL lib citations in our sample would be needed to make the 90th or 50th percentile cutpoints in main Dewey classes and two time periods. Bibliometricians have recommended that percentiles be used for comparisons in *citation* research. “For example,” says Bornmann & Marx (2013: 227), “a value of 90 means that the publication in question is among the 10% most cited publications; the other 90% have achieved fewer citations. A value of 50 indicates the median and therefore an average impact. This way, it is possible to evaluate publications meaningfully and fairly within the same subject category and publication year as a relative scale between 0 (low impact) and 100 (high impact).” In Table 4, lib citations are similarly analyzed. Our

subject categories are overbroad, of course, and we are using six-year rather than single-year periods, but the citation-libcitation parallels are still clear. For instance, mean cutpoints like those at the bottom of the table allow one to give broad advice as to what is distinguished and what is average, such as “If your book sells to 700 or more libraries, you’re in the top 10 percent. A book that sells to 60 or 70 ARL members is doing respectably, but to claim a genuinely large impact, you need over a hundred.”

In Table 4 almost every cutpoint for 2001-2006 is considerably lower than its equivalent for 1990-1995. When the Zuccala-Guns database was created, the books published in the earlier period had had more time to accumulate libcitation counts than those in the later. Also, tighter budgets and costlier serial subscriptions in the later period may have caused libraries to reduce their monographic title purchases. However, Table 4 is meant not as history but simply to imply advisory uses of summary libcitation values.

<i>Libcites</i> 1990-95	<i>Libcites</i> 2001-06	<i>Dewey classes and percentile cutpoints</i>	<i>ARLcites</i> 1990-95	<i>ARLcites</i> 2001-06
		0 Generalia, computing		
784	584	90	110	99
278	194	50	78	45
		1 Philosophy, psychology		
791	671	90	116	107
388	281	50	91	72
		2 Religion		
706	558	90	102	89
284	195	50	56	40
		3 Social sciences		
742	651	90	118	108
320	244	50	87	67
		4 Language		
538	490	90	106	85
195	127	50	65	34
		5 Science		
764	842	90	111	103
328	294	50	85	71
		6 Technology		
735	824	90	111	93
253	231	50	57	44
		7 Arts, recreation		
814	656	90	112	102
357	255	50	85	65
		8 Literature		
844	690	90	114	106
371	242	50	93	67
		9 History, geography		
837	706	90	111	102
296	188	50	77	50
		<i>Means, all classes above</i>		
756	667	90	111	99
307	225	50	77	56

Table 4. Libcitation and ARL libcitation counts as of late 2011 at the 90th and 50th percentiles in main Dewey classes for books published during 1990-1995 or 2001-2006. Above these respective cutpoints, titles are in the top 10% or the top half of their distributions.

Tables 5 through 8 show what famous, average, and little-known books in the main Dewey classes look like. The different distributions underlying them were produced by sorting the sample as described in the captions. The titles in each Dewey class are ordered high, medium, and low, forming another ordinal scale. In general, the *top* title in the group had the highest count in its particular sort (or the highest count after removal of titles used in Tables 2 or 3). The *middle* title had a count at the group median (and may be one of several). The *bottom* title was chosen (by the first author) as typical of those in the tails of distributions, where numerous low-count titles are tied.

The many identically-structured examples allow one to infer some characteristics of books that make for different levels of cultural impact. Ignoring their years of publication, the 20 *high* titles across Dewey classes in Tables 5 and 6 include, in no particular order and with examples:

- Serious popularizations of important subjects (Steven Pinker’s *The language instinct*).
- Broadly useful reference works (the Oxford companions).
- Broadly useful self-help books (Gloria Steinem’s *Revolution from within*).
- Memoirs and biographies of prominent names (Joan Didion’s *The year of magical thinking*, James Gleick’s *Genius* [Richard Feynmann]).
- Books detailing major American historical events (Stephen Ambrose’s *D-Day*).
- Gripping narratives of American scandals (Jon Krakauer’s *Under the banner of heaven*).
- Classic fiction (Flannery O’Connor’s *Complete stories*).

These same categories also fit many of the “super” bestsellers in Tables 2 and 3. Of the top 20 here, four have TV versions (several stories do in O’Connor’s case). Didion’s memoir was adapted for the Broadway stage.

The 20 *medium* or average titles in Tables 5 and 6 take up more specialized topics, and the U.S. emphasis is less apparent or absent altogether. In the 20 *low* titles, the topics are even more specialized, the geographic focus (if present) is highly localized or non-U.S, the temporal focus (if present) is on the distant past, and quite a few are not in English.

Similar observations fit the books with different levels of ARL lib citations in Tables 7 and 8. Again ignoring the time periods, those rising to the top include, in no order and with examples:

- Scholarly analyses of American scandals (Marc Rodwin’s *Money, morals and medicine*).
- Monographs that break new ground (Paul Saint-Amour’s *The copywrights*).
- Introductions to large areas of study (Stephen Gould’s *The structure of evolutionary theory*).
- Distinguished scholarly reference works (Norman Nie’s SPSS manual).

The middle 20 titles once more have narrower topics, while the bottom 20 titles are again very specialized, highly local or non-U.S. in focus, temporally distant, and/or not in English.

<i>Year</i>	<i>Fame</i>	<i>Author</i>	<i>Main Dewey Classes and Titles</i>	<i>Libcites</i>
			<i>0 Generalia, computing</i>	
1991	0	Goodrum	Treasures of the Library of Congress	1994
1993	0	Harris	A patriot press: national politics and the London press in the 1740s	278
1991	0	Turgeon	Inventaire sommaire de la sous-série Affaires de France: 1647-1813	6
			<i>1 Philosophy, psychology</i>	
1992	2	G. Steinem	Revolution from within: a book of self-esteem	2768
1994	1	Vander Waerd	The Socratic movement	388
1993	0	Gill	Superstitions – folk magic in Hull's fishing community	9
			<i>2 Religion</i>	
1993	4	Armstrong	A history of God: the 4000-year quest of Judaism, Christianity, and Islam	2622
1993	1	Schwarzfuchs	A concise history of the rabbinate	284
1992	1	Morgan	The Lambeth Apocalypse, manuscript 209 in Lambeth Palace Library	3
			<i>3 Social sciences</i>	
1991	3	Faludi	Backlash: the undeclared war against American women	3169
1995	0	Gundersen	The environmental promise of democratic deliberation	320
1990	1	Dodds	The silicon forest: high tech in the Portland area, 1945-1986	2
			<i>4 Language</i>	
1995	3	Pinker	The language instinct: how the mind creates language	1985
1993	1	Tobin	Aspect in the English verb: process and result in language	195
1992	1	Hunkin	Gagana Samoa: A Samoan language coursebook	5
			<i>5 Science</i>	
1992	2	Gleick	Genius: the life and science of Richard Feynman	2538
1990	0	Potter	Sets: an introduction	328
1994	0	Walde	The Mortlach phase	3
			<i>6 Technology</i>	
1993	4	B. Moyers	Healing and the mind	3204
1994	1	Cannon	Dye plants and dyeing	253
1991	2	Bombardier Inc.	Ski-doo, as long as there's winter—	3
			<i>7 Arts, recreation</i>	
1993	2	R. Hughes	Culture of complaint: the fraying of America	2133
1991	0	Essick	William Blake's commercial book illustrations: a catalogue and study of the plates engraved by Blake after designs by other artists	357
1994	1	Bon / Place	François Place, illustrateur	3
			<i>8 Literature</i>	
1994	4	F. O'Connor	The complete stories	2953
1993	1	Terdiman	Present past: modernity and the memory crisis	371
1991	0	Lee	Daurel e Beton	8
			<i>9 History, geography</i>	
1994	2	Ambrose	D-Day: June 6, 1944: the climactic battle of World War II	3406
1990	1	Yarak	Asante and the Dutch, 1744-1873	296
1991	0	Yeebo	Ghana: the struggle for popular power: Rawlings, saviour or demagogue	3

Table 5. Titles from 1990-1995 with lib citations at (or near) the top, middle, and bottom of the count distributions in main Dewey classes. Fame scale: 4, Wikipedia title and author entries, plus TV version; 3, Wikipedia title and author entries; 2, Wikipedia author entry; 1, Wikipedia citation; 0, no Wikipedia data.

<i>Year</i>	<i>Fame</i>	<i>Author</i>	<i>Main Dewey Classes and Titles</i>	<i>Libcites</i>
			<i>0 Generalia, computing</i>	
2002	0	Knowles	The Oxford dictionary of modern quotations	2592
2002	2	Hamilton	Refiguring the archive	194
2006	1	Bergmann	Medienmenschen: wie man wirklichkeit inszeniert	3
			<i>1 Philosophy, psychology</i>	
2005	3	Honderich, ed.	Oxford companion to philosophy	2498
2005	0	Carone	Plato's cosmology and its ethical dimensions	281
2005	0	York	Respect for the world: universal ethics and the morality of terraforming	3
			<i>2 Religion</i>	
2003	3	Krakauer	Under the banner of heaven: a story of violent faith	2931
2001	1	Murphy	Jesuit slaveholding in Maryland: 1717-1838	195
2001	2	Roth	História dos marranos: os judeus secretos da Península Ibérica	2
			<i>3 Social sciences</i>	
2006	3	J. Carter	Our endangered values: America's moral crisis	3380
2004	1	Edkins	Sovereign lives: power in global politics	244
2006	0	Wilson	The beat: policing a Victorian city	2
			<i>4 Language</i>	
2002	4	McCrum	The story of English	3433
2005	1	Göksel	Turkish: a comprehensive grammar	127
2001	0	Caprini	Nomi propri	7
			<i>5 Science</i>	
2004	3	Bryson	A short history of nearly everything	3284
2001	1	Scarth	Volcanoes of Europe	294
2003	0	Blanchard	L'art populaire dans le briançonnais: les cadrans solaires	3
			<i>6 Technology</i>	
2003	3	Agatston	The South Beach diet	3088
2005	1	Cunfer	On the Great Plains: agriculture and environment	231
2001	0	Giannetti	Trattato di dermatologia	2
			<i>7 Arts, recreation</i>	
2004	2	Bordman	The Oxford companion to American theatre	2140
2003	1	Williams	Cricket and England: a cultural and social history of the inter-war years	255
2006	0	Moller	Technicolor dreamin': the 1960s rainbow and beyond	2
			<i>8 Literature</i>	
2005	4	Didion	The year of magical thinking	3337
2002	0	Worman	The cast of character: style in Greek literature	242
2006	0	Faranda	Via seminario, 19	3
			<i>9 History, geography</i>	
2001	2	J. Carter	An hour before daylight: memories of a rural boyhood	3383
2003	0	Campbell	English public opinion and the American Civil War	188
2006	0	Cust	Cases in the court of chivalry, 1634-1640	2

Table 6. Titles from 2001-2006 with lib citations at (or near) the top, middle, and bottom of the count distributions in main Dewey classes. Fame scale: 4, Wikipedia title and author entries, plus TV or (Didion) Broadway version; 3, Wikipedia title and author entries; 2, Wikipedia author entry; 1, Wikipedia citation; 0, no Wikipedia data.

<i>Year</i>	<i>Fame</i>	<i>Author</i>	<i>Main Dewey Classes and Titles</i>	<i>ARLcites</i>
			<i>0 Generalia, computing</i>	
1991	3	Nie	Statistical package for the social sciences	137
1992	2	Tsonis	Chaos: from theory to applications	79
1992	2	El-Abbadi	Vie et destin de l'ancienne Bibliothèque d'Alexandrie	2
			<i>1 Philosophy, psychology</i>	
1993	1	Rodwin	Medicine, money and morals: physicians' conflicts of interest	161
1993	2	Csikszentmihalyi	The evolving self: a psychology for the third millennium	91
1993	2	Guha	A construction of humanism in colonial India	1
			<i>2 Religion</i>	
1995	2	Tabor	Why Waco?: cults and the battle for religious freedom in America	137
1994	2	Hastings	The church in Africa: 1450-1950	56
1990	0	Hinnebusch	Brève histoire de l'ordre dominicain	1
			<i>3 Social sciences</i>	
1990	2	Young	Justice and the politics of difference	165
1992	0	Hall	Common threads: a parade of American clothing	87
1990	0	Guildford	Public school reform and the Halifax middle class, 1850-1870	1
			<i>4 Language</i>	
1990	2	Bickerton	Language and species	124
1995	2	Lasersohn	Plurality, conjunction and events	65
1993	2	de Urbina	Diccionario manual griego-español	1
			<i>5 Science</i>	
1992	1	C. Morris, ed.	Academic Press dictionary of science and technology	146
1994	2	Goodwin	How the leopard changed its spots: The evolution of complexity	85
1991	1	Benson	Rare bushland plants of Western Sydney	1
			<i>6 Technology</i>	
1993	0	Kiple	The Cambridge world history of human disease	153
1993	0	Durling	A dictionary of medical terms in Galen	57
1994	0	Heikell	Greek waters pilot	1
			<i>7 Arts, recreation</i>	
1990	1	Reid	The Oxford guide to classical mythology in the arts, 1300-1990s	132
1990	2	Jones	Fake? the art of deception	85
1993	2	Simmons	Image of the train: the Victorian era	1
			<i>8 Literature</i>	
1991	2	T. Grey	The Wallace Stevens case: law and the practice of poetry	155
1993	0	Gera	Xenophon's Cyropaedia: style, genre, and literary technique	93
1995	0	Neumann	Schlemihl's travels: public lecture delivered on 28 September 1994	3
			<i>9 History, geography</i>	
1991	2	Lazarus	Black Hills / white justice: the Sioux Nation versus the United States, 1775 to the present	155
1994	0	Marley	Pirates and privateers of the Americas	77
1995	0	Petrenko	The archaeological monuments of Grobina	1

Table 7. Titles from 1990-1995 with ARL lib citations at (or near) the top, middle, and bottom of the distributions in main Dewey classes. Fame scale: 3, Wikipedia title and author entries; 2, Wikipedia author entry; 1, Wikipedia citation; 0, no Wikipedia data.

<i>Year</i>	<i>Fame</i>	<i>Author</i>	<i>Main Dewey Classes and Titles</i>	<i>ARLcites</i>
			<i>0 Generalia, computing</i>	
2001	3	N. Baker	Double fold: libraries and the assault on paper	155
2002	2	D. Wilson	The British Museum: a history	45
2003	0	Witt	The Carnegie libraries of Iowa	2
			<i>1 Philosophy, psychology</i>	
2001	2	Beauchamp	The principles of biomedical ethics	160
2005	0	Hugman	New approaches in ethics for the caring professions	72
2003	0	Brunori	Traumi di guerra. un'esperienza psicoanalitica in Bosnia-Erzegovina	2
			<i>2 Religion</i>	
2004	2	Pelikan	Interpreting the Bible and the Constitution	141
2004	2	Bottum, ed.	The Pius war: responses to the critics of Pius XII	40
2004	0	Tamas	The wrath of gods: esoteric and occult in the modern world	3
			<i>3 Social sciences</i>	
2002	2	Banner	The death penalty: an American history	164
2003	2	Moxham	Tea: addiction, exploitation, and empire	67
2002	0	Otter	Solomon Islands human development report 2002: building a nation	2
			<i>4 Language</i>	
2002	2	J. Hall, ed.	Dictionary of American regional English	142
2003	0	Corbett	An intercultural approach to English language teaching	34
2004	0	Shnukal	Dictionary of Torres Strait Creole	1
			<i>5 Science</i>	
2002	3	Gould	The structure of evolutionary theory	133
2003	2	Murray	Mathematical biology	71
2006	1	Van Landuyt	Atlas van de flora van Vlaanderen en het Brussels Gewest	2
			<i>6 Technology</i>	
2003	1	Babcock	Women don't ask: negotiation and the gender divide	155
2006	1	Lim	Tilapia: biology, culture, and nutrition	44
2006	1	Rémy	Conseiller en agriculture	2
			<i>7 Arts, recreation</i>	
2003	2	Rapping	Law and justice as seen on TV	140
2006	1	Futrell	The Roman games. a sourcebook	65
2004	0	Eisenberg	FIFA 1904-2004 : le siècle du football	3
			<i>8 Literature</i>	
2003	0	Saint-Amour	The copywrights: intellectual property and the literary imagination	138
2005	1	Johnson	The fountainheads: Wright, Rand, the FBI and Hollywood	67
2006	1	Guaragnella	Studi di letteratura Italiana: per Vitorio Masiello	1
			<i>9 History, geography</i>	
2005	3	Greenberg, ed.	The torture papers: the road to Abu Ghraib	143
2003	2	Cosh	Edinburgh: the golden age	50
2003	2	Brooks	Church, state and access to resources in early Anglo-Saxon England	1

Table 8. Titles from 2001-2006 with ARL lib citations at (or near) the top, middle, and bottom of the distributions in main Dewey classes. Fame scale: 3, Wikipedia title and author entries; 2, Wikipedia author entry; 1, Wikipedia citation; 0, no Wikipedia data. *The Dictionary of American regional English*, scored 2, has a Wikipedia entry for title, not author.

	<i>Total lib citations</i>					<i>ARL lib citations</i>			
<i>Wikipedia fame scale</i>	1 Low	2 Med	3 High	<i>Totals</i>	<i>Wikipedia fame scale</i>	1 Low	2 Med	3 High	<i>Totals</i>
3 Title/author articles	0	0	12	12	3 Title/author articles	0	0	4	4
2 Author article only	2	1	6	9	2 Author article only	5	11	10	26
1 Cited only	5	12	0	17	1 Cited only	3	3	4	10
0 None	13	7	2	22	0 None	12	6	2	20
<i>Totals</i>	20	20	20	60	<i>Totals</i>	20	20	20	60

Tables 9a and 9b. Tests of association between the Wikipedia fame scale and total lib citation counts (Tables 5-6) or ARL lib citation counts (Tables 7-8). In Table 9a, titles scored 3 on the fame scale include five that actually scored 4. This equalizes the tables in size without affecting test results. For Table 9a, Gamma = 0.77, $p < .000$. For Table 9b, Gamma = 0.60, $p < .000$.

Given each title's high-to-low lib citation counts in its Dewey class and high-to-low placement on the fame scale, we have two variables that can be tested for strength of association. The two variables are symmetric (i.e., neither causes the other); both simply operationalize what we have called fame. A standard measure of association for symmetric ordinal data with many tied values is Goodman and Kruskal's gamma (Sirkin, 2006: 367), which ranges from -1 to $+1$, with 0 indicating no relationship. Tables 9a and 9b display the distributions and the test results.

These results support the argument that, just as Wikipedia entries capture the relative fame of books, so do lib citations. The gamma of 0.77 in Table 9a indicates a very strong relationship; the gamma of 0.60 in Table 9b, a substantial one. The probabilities that these relationships occurred by chance are extremely low ($p < .000$). All alternative SPSS measures of association for these two tables are also substantial, with $p < .000$.

To gloss Table 9a, most titles with *low* lib citation counts are in the *bottom* two rows of the Wikipedia variable (13+5); most *medium* titles are in the *middle* two Wikipedia rows (12+1), and most *high* titles are in the *top* two Wikipedia rows (6+12). This is akin to a scatterplot for two ratio-level variables that shows a strong, direct relationship between their low, medium, and high values. Table 9b may be read in the same way. There, however, the majorities of titles in the bottom, middle and top Wikipedia rows are smaller—and the “off-pattern” titles more numerous—which somewhat lessens the relationship.

Overall, these results do not suggest that lib citation counts have been manipulated or that authors and publishers could easily manipulate them. The latter claim, by Hammarfelt (2016: 122), requires demonstration, especially with regards to WorldCat, where decentralized acquisitions produce counts that range over four orders of magnitude. It is not at all clear how this system could be gamed for unfair advantage, or how anyone who tried it could escape being found out.

Conclusion

Having shown particular books at various levels of fame and also characterized the kinds of books that occupy these different levels, we return to the notion of cultural impact. If this is equated with marked social benefit, such as stimulating new legislation or reforming some evil, lib citations may seem a very remote measure. On the contrary, books associated with important changes of any sort—certainly major social ones—are likely to have correspondingly high lib citation counts, whether relative or absolute.

Although we have stressed lib citations as an altmetric for books in the humanities and social sciences, Tables 5 through 8 also display them for several advanced monographs in science and technology, suggesting they could complement citations there as well. The *content* of these latter fields is different, but the mechanisms of fame in them probably are not. In any case, the idea could be tested with sci-tech books in a future study.

Table 2 displays books whose cultural impact is undeniable. At minimum, the books in Tables 2 and 3 have entered into national or international dialogues, and a “reception story” could be assembled for any of them. Moreover, if this can be done for the titles in Tables 2 and 3, it can also be done for titles in the other tables.

Critics of bibliometric measures often insinuate that the evaluators who use them will look only at the potentially misleading numbers. The numbers, however, are mere indicators that always point to a story. Authors of books may be uniquely qualified to tell their stories in this sense and can add those accounts in evaluations, just as citees can elaborate on the citations their work has received. This very point is extensively made for research in the arts and humanities by Thelwall & Delgado (2015). However, their solution is for stories (which they call “data”) to drive out metrics—in their case, citation measures—completely. They do not consider that lib citations could be a metric that supplements and reinforces scholars’ stories about such research.

That said, these stories are cases—arguments—that a book has had an impact, and, just as happens with citations, some cases will be better than others. Thus, for authors of books little held by libraries, lib citation counts will not seem an attractive metric. But for authors of books whose sales to libraries are comparatively high, they are one more proof of achievement to add to the record.

Acknowledgements

We thank the Elsevier Bibliometrics Research Programme and OCLC for the data used in the research database. We also thank Raf Guns and Roberto Cornacchia for helping to develop it, and Maurits van Bellen and Robert Iepsma for their data cleaning and standardization work.

References

- Adler, M. (2009). Transcending library catalogs: A comparative study of controlled terms in Library of Congress Subject Headings and user-generated tags in LibraryThing for transgender books. *Journal of Web Librarianship*, 3(4), 309-331.
- Ajiferuke, I., & Wolfram, D. (2010). Citer analysis as a measure of research impact: Library and information science as a case study, *Scientometrics*, 83(3), 623-638.
- Altmetric. (2017). What are altmetrics? <https://www.altmetric.com/about-altmetrics/what-are-altmetrics/>
- Bornmann, L., & Marx, W. (2013). How good is research really? *EMBO Reports*, 14(3), 226-230.
- Denton, W. (2012). On dentographs, a new method of visualizing library collections. *Code4Lib Journal*, 16, 171-179. <http://journal.code4lib.org/articles/6300>
- Gorraiz, J., Purnell, P. J., & Glänzel, W. (2013). Opportunities for and limitations of the Book Citation Index. *Journal of the American Society for Information Science and Technology*, 64(7), 1388-1398.
- Halevi, G., Nicolas, B., Bar-Ilan, J. (2016). The complexity of measuring the impact of books. *Publishing Research Quarterly*, 32(3), 187-200.
- Hammarfelt, B. (2016). Beyond coverage: Toward a bibliometrics for the humanities. In Ochsner, M., Hug, S. E., & Daniel, H.-D. (Eds.) *Research assessment in the humanities: Towards criteria and procedures*. Springer Open. 115-131. https://link.springer.com/chapter/10.1007%2F978-3-319-29016-4_10
- Kousha, K., & Thelwall, M. (2015). Alternative metrics for book impact assessment: Can Choice reviews be a useful source? *Proceedings of the 15th International Society for Scientometrics and Informetrics Conference*, v. 1, 59-70.
- Kousha, K., & Thelwall, M. (2016). Can Amazon.com reviews help to assess the wider impacts of books? *Journal of the Association for Information Science and Technology*, 67(3), 566-581.
- Kousha, K., & Thelwall, M. (2017). Are Wikipedia citations important evidence of the impact of scholarly articles and books? *Journal of the Association for Information Science and Technology*, 68(3), 762-779.
- Kousha, K., Thelwall, M., & Abdoli, M. (2017). Goodreads reviews to assess the wider impacts of books. *Journal of the Association for Information Science and Technology*, 68(8), 2004-2016.
- Kousha, K., Thelwall, M., & Rezaie, S. (2011). Assessing the citation impact of books: The role of Google Books, Google Scholar, and Scopus. *Journal of the American Society for Information Science and Technology*, 62(11), 2147-2164.
- Linmans, A. J. M. (2010). Why with bibliometrics the humanities does not need to be the weakest link. *Scientometrics*, 83(2), 337-354.
- Ochsner, M., Hug, S., & Gallerson, I. (2017). The future of research assessment in the humanities: Bottom-up assessment procedures. *Palgrave Communications*, 3. <http://dx.doi.org/10.1057/palcomms.2017.20>
- OCLC. (2004). About the OCLC top 1000. <http://www.oclc.org/research/top1000/about.html> Complete list. <http://www.oclc.org/research/top1000/complete-2004.html>
- Plum Analytics. (2017). Usage metrics. <https://plumanalytics.com/learn/about-metrics/usage-metrics/>
- Sirkin, R.M. (2006). *Statistics for the social sciences*, 3d ed. Thousand Oaks, CA: Sage.

- Thelwall, M. (2017). Web indicators for research evaluation: A practical guide. *Synthesis Lectures on Information Concepts, Retrieval, and Services*, 52. Morgan & Claypool.
- Thelwall, M., & Delgado, M. M. (2015). Arts and humanities research evaluation: No metrics please, just data. *Journal of Documentation*, 71(4), 817-833.
- Torres-Salinas, D., Gumpenberger, C., & Gorraiz, J. (2017). PlumX as a potential tool to assess the macroscopic multidimensional impact of books. *Frontiers in Research Metrics and Analytics* 2:5.
<https://www.frontiersin.org/articles/10.3389/frma.2017.00005/full>
- Torres-Salinas, D., & Moed, H. F. (2009). Library catalog analysis as a tool in studies of social sciences and humanities: An exploratory study of published book titles in economics. *Journal of Informetrics*, 3(1), 9-26.
- Waltman, L. (2016). A review of the literature on citation impact indicators. *Journal of Informetrics*, 10(2), 365-391.
- White, H. D. (2008). Better than brief tests: Coverage power tests of collection strength. *College and Research Libraries*, 69(2), 155-174.
- White, H. D., Boell, S. K., Yu, H., Davis, M., Wilson, C. S., & Cole, F. T. H. (2009). Libcitations: A measure for comparative assessment of book publications in the humanities and social sciences. *Journal of the American Society for Information Science and Technology*, 60(6), 1083-1096.
- Zhou, Q., Zhang, C., Zhao, S. X., & Chen, B. (2016). Measuring book impact based on the multi-granularity online review mining. *Scientometrics*, 107(3), 1435-1455.
- Zuccala, A., & Guns, R. (2013). Comparing book citations in humanities journals to library holdings: Scholarly use versus 'perceived cultural benefit' (RIP). *Proceedings of the 14th International Society for Scientometrics and Informetrics Conference*, v.1, 353-360.
- Zuccala, A., Guns, R., Cornacchia, R., Bod, R. (2015). Can we rank scholarly book publishers? A bibliometric experiment with the field of history. *Journal of the Association for Information Science and Technology*, 66(7), 1333-1347.
- Zuccala, A., Someren, M., & Bellen, M. (2014). A machine-learning approach to coding book reviews as quality indicators: Toward a theory of megacitation. *Journal of the Association for Information Science and Technology*, 65(11), 2248-2260.
- Zuccala, A. A., & White, H. D. (2015). Correlating libcitations and citations in the humanities with WorldCat and Scopus data. *Proceedings of the 15th International Society for Scientometrics and Informetrics Conference*, Part 2, 305-316.